LOYOLA COLLEGE (AUTONOMOUS) CHENNAI – 600 034

M.Sc. DEGREE EXAMINATION – **STATISTICS**

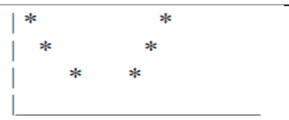
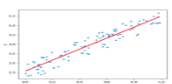
FIRST SEMESTER – **NOVEMBER 2024**

PST1MC02 – APPLIED REGRESSION ANALYSIS

Date: 11-11-2024 Dept. No. Max. : 100 Marks

Time: 01:00 pm-04:00 pm

	SECTION A – K1 (CO1)
	Answer ALL the questions **(5 x 1 = 5)**
1	**MCQ**
a)	The value of R^2 lies between: i) -1 and 0 ii) -1 and 1 iii) 0 and 1 iv) None of the above
b)	What is the primary advantage of polynomial regression? i) Interpretability ii) handling non-linearity iii) Reducing Multicollinearity iv) Improving the model
c)	Which one of the following is not a variance stabilizing transformation? i) Log Transformation ii) Square root Transformation iii) Arcsin Transformation iv) Min-Max scaling
d)	In a model $Y = f(x_1, x_2, x_3)\, if\ r_{12} = r_{23} = r_{13} = 1$, then it is said to be i) Multicollinearity ii) perfect multicollinearity iii) orthogonality iv) None of these
e)	One of the most widely used methodologies for the analysis of time series data is i) ARIMA ii) AR iii) MA iv) ARMA

	SECTION A – K2 (CO1)	
	Answer ALL the questions	**(5 x 1 = 5)**
2	**Match the following**	
a)	SLM	Weighted Least Square method
b)	Polinomial Regression	VIF
c)	Transformation	ARIMA
d)	Multicollinearity	
e)	Time Series Forecasting	

	SECTION B – K3 (CO2)
	Answer any THREE of the following **(3 x 10 = 30)**
3	Obtain the Least squares estimator of Multiple Linear Regression coefficients
4	Describe Cubic Spline fitting
5	The residuals from a model arranged in time-order had the following signs + + + + − − − − − − − + − − + + − − − + + + − − + + − − − +. Carry out the Runs Test for randomness of the sequence.
6	Explain Weighted least square method.
7	Describe the methods of detecting Non-linearity.

	SECTION C – K4 (CO3)
	Answer any TWO of the following (2 x 12.5 = 25)
8	In multiple regression modeling, give rough graphical illustrations of the different possible scenarios when the residuals are plotted against the predicted values. Describe how these are used for model modification.
9	Explain the use of a dummy variable in regression analysis.
10	List down four sources of multicollinearity with illustrations.
11	In building a model, the OLS residuals are : 0.6,1.9,-1.8,-2.7,-2.9,1.4,3.3,0.3,0.8,2.3,-1.4,-1.1. Carry out the Durbin-Watson test at 5% significance level and draw your conclusion on the presence of autocorrelation [It is given that d_L =1.45 , d_U = 1.65]
	SECTION D – K5 (CO4)
	Answer any ONE of the following (1 x 15 = 15)
12	Carry out the 'Forward Model Building' Process to build a model with four regressors given the following information on SSRes for different subset models with a sample of size 20. Use a significance of 5%: SStotal = 237.629, SSRes(X_1) = 169.698, SSRes(X_2) = 79.303, SSRes(X_3) = 110.749, SSRes(X_4) = 77.340, SSRes(X_1, X_2) = 36.351, SSRes(X_1, X_3) = 107.368, SSRes(X_1, X_4) = 15.377, SSRes(X_2, X_3) = 5.068, SSRes(X_2, X_4) = 76.027, SSRes(X_3, X_4) = 6.542, SSRes(X_1, X_2, X_3) = 4.211, SSRes(X_1, X_2, X_4) = 6.458, SSRes(X_1, X_3, X_4) = 4.449, SSRes(X_2, X_3, X_4) = 4.198, SSRes(X_1, X_2, X_3, X_4,) = 4.188
13	Explain the different types of transformations used in modelling.
	SECTION E – K6 (CO5)
	Answer any ONE of the following (1 x 20 = 20)
14	a) Explain any two uses of Residual analysis. (6) b) Write down the multiple linear model equation and intercept of the model coefficients based on the following information (14) $(X'X)^{-1} = \begin{bmatrix} 2 & 6 & -2 \\ 6 & 7 & -4 \\ -2 & -4 & 3 \end{bmatrix}, X'Y = \begin{bmatrix} 3 \\ 7 \\ 10 \end{bmatrix}$ Compute (ii) Overall significance of the model fit (iii) Confidence interval for slope (iv) interpret R–square
15	Explain the various methods of diagnosing multicollinearity and suggest the methods for removing it.

$$$$$$$$$$$$$$$